Reply to Dr. Katz by Dr. Barreto, January 18, 2022

The report submitted by Dr. Jonathan Katz after the conclusion of expert testimony on January 14, 2022 has no relevance whatsoever to understanding Voting Rights Act issues, or voting patterns in Pennsylvania.

Preliminarily, it is important to underscore what is not in Dr. Katz's report. Dr. Katz does not opine that there is a lack of cohesiveness in minority voting patterns in Pennsylvania or that there is no racially polarized voting in Pennsylvania. In short, he makes no effort to actually use statistical methods to contribute to our understanding of voting patterns in Pennsylvania. His essential premise—that Ecological Inference or EI is not a proper basis for evaluating Latino political cohesiveness in Pennsylvania—is not valid. Ecological inference has been extensively published as a reliable method in political science journals, and has been regularly used, in fact, required, by state and federal courts in adjudicating voting rights lawsuits. That Dr. Katz finds it challenging to use widely accepted scientific methodology to identify and understand racially polarized voting in Pennsylvania is both suspect and unremarkable.

It is also noteworthy that Dr. Katz does not dispute that Black voters in Pennsylvania are very cohesive and clearly meet the second *Gingles* criteria of minority cohesiveness. Further, Dr. Katz does not dispute that White voters in Pennsylvania block vote against minority candidates of choice, meeting the third *Gingles* criteria of bloc-voting. To be clear, there is no meaningful analysis or data in Dr. Katz's report concerning application of the *Gingles* test in Pennsylania.

Dr. Katz's factual assertions with respect to Latinos in Pennsylvania are not accurate. Dr. Katz claims that "there are no homogenous Latino/Hispanic voting precincts in the state." This is not accurate. In fact, there are 23 precincts in Pennsylvania which are over 80% Latino, with an additional 51 precincts which are greater than 70% Latino. Indeed across the state there are 213 precincts which are majority-Latino, and there are a total of 283 precincts in which Latinos are the single largest racial group (i.e. precincts where Latinos make up less than 50% but outnumber Whites, Blacks and all other groups.). I have attached a list of all 283 such precincts with the State House 2020 votes for Democratic and Republican candidates as an appendix.

These heavily Latino precincts provide clear data to support the conclusion that Latino voters in Pennsylvania are cohesive and support Democrats. That data is summarized in Table 1 below.

Table 1: Average State House Democratic Vote 2020 in Heavily Latino Precincts

% Latino Range	Precincts	Avg. Dem vote	Std. Dev	Min	Max
80 – 100%	N=23	97.3%	.0712	.7704	1.00
70 – 100%	N=74	91.8%	.1256	.4588	1.00
50 – 100%	N=213	89.9%	.1613	.3346	1.00
Plurality	N=283	90.3%	.1562	.3356	1.00

In my presentation on January 14, 2022 to the LRC, I reported an estimated Latino vote ranging from 74% to 82% for State House legislative elections across different regions of the state. Dr.

Katz provides no contrary opinions concerning the level of Latino cohesion, instead just opining that ecological inference is complicated and contains uncertainty. It is true that all of social science empirical analysis is complex and there is not perfect certainty. Our job as trained methodologists and social scientists is to use the best known tools and models to reduce uncertainty and provide best available estimates from which we can draw inferences and conclusions. Ecological inference models to determine racial and ethnic voting patterns are a widely accepted and published methodology in political science.

Dr. Katz criticizes the ecological inference method in his report, but he has not published any peer-reviewed academic articles finding that ecological inference is unreliable. His opposition to ecological inference runs counter to an abundance of social science published research[1] that support ecological inference as an appropriate tool for estimating racial and ethnic voting patterns and its accepted use by voting rights experts in evaluating whether or not racially polarized voting exists.

Other critiques made by Dr. Katz are also baseless. He includes a section claiming that King's EI model is generally wrong and does not work well in elections with multiple racial groups or multiple candidates (Sections 3.3 and 3.4). All analyses in my report however are based on two-candidate elections. Dr. Katz criticizes King (1997) and advocates for the use of Rosen et al. (2001), yet the full analysis that I conducted and reported is based on eiCompare software package which uses *both* King and Rosen, and allows the analyst to *compare* how they perform. In published research,[2] we have twice demonstrated that the King and Rosen methods are highly correlated with one another and both provide accurate results. In this case, my analysis relies on both the King and Rosen approaches to ecological inference within the eiCompare software package.

---

[1] Grofman, Bernard. 1991. "Statistics without substance: A critique of Freedman et al." Evaluation Review, 15: 746-769; Lichtman, Alan. 1991. "Passing the Test." Evaluation Review. 15, 770-799.; Tanner, Martin. 1996. Tools for statistical inference: methods for the exploration of posterior distributions and likelihood functions, 3rd Ed., Springer, New York; King, Gary. 1997. A Solution to the Ecological Inference Problem. Princeton University Press, King, Gary, Ori Rosen and Martin Tanner. 1999. "Binomial-Beta hierarchical models for ecological inference" Sociological Methods and Research, 28: 61-90; King, Gary. 1999. "The Future of Ecological Inference Research: A Comment on Freedman et al." Journal of the American Statistical Association Vol. 94, No. 445 (Mar., 1999), pp. 352-355; Rosen, Ori, Wenxin Jiang, Gary King, Martin Tanner. 2001. "Bayesian and frequentist inference for ecological inference: the RxC case" Statistica Neerlandica. 55:2; Grofman, Bernard and Samuel Merrill. 2004. "Ecological Regression and Ecological Inference." In Gary King et al., eds. Ecological Inference: New Methodological Strategies. Cambridge University Press.; Barreto, Matt 2007. "Si Se Puede! Latino Candidates and the Mobilization of Latino Voters." American Political Science Review. 101 (August); Grofman, Bernard and Matt Barreto. 2009. "A Reply to Zax's (2002) Critique of Grofman and Migalski (1988): Double Equation Approaches to Ecological Inferences" Sociological Methods and Research. 37 (May); Collingwood, Loren, Kassra Oskooii, Sergio Garcia-Rios, and Matt Barreto. 2016. "eiCompare: Comparing Ecological Inference Estimates across EI and EI: RxC." The R Journal. 8:2; Imai, Kosuke and Kabir Khanna. 2016. "Improving Ecological Inference by Predicting Individual Ethnicity from Voter Registration Records" Political Analysis; Barreto, Matt, Loren Collingwood, Sergio Garcia-Rios and Kassra Oskooii. 2019. "Estimating Candidate Support: Comparing Iterative EI and EI-RxC Methods" Sociological Methods and Research. 48(4).

[2] Collingwood, Loren, Kassra Oskooii, Sergio Garcia-Rios, and Matt Barreto. 2016. "eiCompare: Comparing Ecological Inference Estimates across EI and EI: RxC." The R Journal. 8:2; Barreto, Matt, Loren Collingwood, Sergio Garcia-Rios and Kassra Oskooii. 2019. "Estimating Candidate Support: Comparing Iterative EI and EI-RxC Methods" Sociological Methods and Research. 48(4).

In Section 4 of his report Dr. Katz suggests that ecological inference is not appropriate for Pennsylvania due to lack of homogeneous precincts, but the only evidence he relies on is data from California. He provides no relevant evidence in this section that my conclusions concerning Pennsylvania are inaccurate.

Curiously, to support his argument that ecological inference is not valid for understanding Latino voting patterns in Pennsylvania, Dr. Katz relies primarily on Latino party registration rates in Kern County, California from years ago. Even here, his analysis of Latino party registration in Kern County, California is deeply flawed and did not properly consider voters who had no party registration which is why his model over-estimated Democratic registration. In the case Dr. Katz cited, *Luna v. County of Kern*, 291 F. Supp. 3d 1088 (E.D. Cal. 2018), Dr. Katz's critique of ecological inference was rejected by the court. With respect to "Dr. Katz's critiques," the decision states: "[T]he court is unpersuaded that these criticisms preclude plaintiffs from demonstrating Latino political cohesiveness by a preponderance of the evidence." The court found "no basis to conclude that there is some minimum number of homogenous precincts required before [ecological regression] and [ecological inference] analysis have any probative value" in a VRA case. The court noted that Dr. Katz himself admitted that the political scientist who developed ecological inference (Gary King) "indicated no bright line percentage of homogenous precincts is necessary in order for ecological inference estimates to be reliable." The court in *Luna* further noted that, in addition to the lack of support for Dr. Katz's position "in the field of statistics, numerous cases finding racial polarization have relied on statistical analyses that did not include HPA [homogenous precinct analysis] and made no mention of homogenous precincts whatsoever." Finally, the court found that "Dr. Katz's insistence on 'sufficient' homogenous precincts is undercut by his own work in previous cases, where he performed [ecological regression] and [ecological inference] analyses without any reference to the number of homogenous precincts in the relevant jurisdiction." The *Luna* court held that Dr. Katz's critique did "not raise a doubt sufficient to refute" plaintiffs' expert's conclusion that racial polarization existed.

Given that a federal judge so soundly dismissed Dr. Katz's theory concerning homogenous precincts, the Commission should question why such a debunked theory was offered at the very last moment. The late submission suggests that proponents of Dr. Katz's report held it until the 11th hour to shield both Dr. Katz and his report from fair examination and scrutiny.
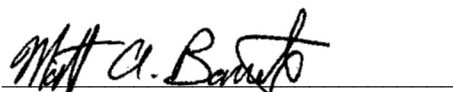
Dr. Katz claims on page 10 of his report that my analysis "focuses almost exclusively on statewide offices." However in every single instance I provide vote pattern results and estimates for State House elections, which would be considered the endogenous elections. To provide the Commission with *more* information and evidence, I also include voting patterns in additional elections. As Dr. Warshaw stated in his testimony on January 14, 2022, it is a longstanding conclusion in Political Science that voting in State Legislative elections is highly correlated with Presidential elections and elections for other major partisan statewide offices. If Dr. Katz's argument were correct, we would not observe a strong correlation between statewide elections and district elections, but the correlation is undeniable. Excluding elections in which a candidate ran unopposed, the correlation between Democratic vote for State House and Democratic vote for Attorney General at the precinct level is 0.9448 in 2020. The correlation between

Democratic vote for State House and Democratic vote for U.S. President at the precinct level is 0.9233 in 2020. Thus, as decades of political science literature suggest, these elections are highly consistent.

There was a question at the hearing concerning consideration of primary elections, but neither court precedent nor peer-reviewed political science literature require an evaluation of primary election results to draw conclusions about racially polarized voting. Indeed, in this instance, we are interested in whether or not Whites, Blacks, and Latinos vote for the same or different candidates to represent them and the most probative elections here are general elections where voters are choosing which candidates to send to the state Capitol. While primary election data may be instructive in case specific situations where minorities do not vote in coalition, there is strong evidence in Pennsylvania that minorities do vote in coalition. Moreover, as the Chairman noted, the preliminary plan creates a number of open seats which will provide opportunities for minority candidates in primary elections.

To be clear, the precinct scatterplots are presented because they are illustrative examples of voting patterns that are clear, concise and easy to interpret. However, the full ecological inference models are based on the eiCompare which estimates models for Whites, minorities-combined, Blacks, and Latinos, using both ecological methods advocated by King and Rosen et al.

In summary, the report by Dr. Katz concedes much and does not offer any evidence or data on voting patterns in Pennsylvania. Dr. Katz only raises generic critiques of ecological inference which have been debunked by courts and in the social science literature. Read in full, his report is non-responsive and offers nothing that in any way detracts from the well-supported conclusions and opinions in my report.

Matt A. Barreto
January 18, 2022
Agoura Hills, California